

Denkanstöße zum Einsatz von ADM-Systemen in der öffentlichen Verwaltung

Arbeitspapier aus dem BMBF-geförderten Forschungsprojekt „FairAndGood ADM“
(Teilprojekt Politikwissenschaft)

*Maryam Haeri, Kathrin Hartmann, Pascal König, Tobias Krafft, Jürgen Sirsch,
Karen Joisten, Georg Wenzelburger & Katharina Zweig*

Algorithmische Entscheidungssysteme (Algorithmic Decision Making (ADM)-Systeme) halten zunehmend Einzug in die staatliche Verwaltung. In Deutschland sind die Anwendungsfälle noch vergleichsweise überschaubar: So wendet etwa die Polizei in manchen Bundesländern ADMs zur Vorhersage von Kriminalitätswahrscheinlichkeiten an (Egbert and Krasmann 2019) und nutzt diese in der vorausschauenden Polizeiarbeit („Predictive Policing“). Und in der Finanzverwaltung prüfen algorithmische Systeme die Plausibilität von Steuererklärungen und erlauben so, auffällige Erklärungen herauszufiltern, die dann tiefergehend bearbeitet werden (Martini 2019: 17). Dennoch: In anderen Staaten – und insbesondere in den USA – werden ADMs deutlich umfangreicher in der öffentlichen Verwaltung eingesetzt (Wissenschaftliche Dienste des Bundestages 2019), sodass aus den dort gemachten Erfahrungen auch Lehren für mögliche weitere Anwendungen in Deutschland gezogen werden können.

In diesem kurzen Diskussionspapier möchten wir erste vorläufige Schlussfolgerungen aus drei explorativen Fallstudien in Form von fünf Denkanstößen pointiert zusammenfassen und damit zur politischen und gesellschaftlichen Diskussion beitragen. Die Fallstudien, welche die empirische Basis für die Ableitung der Denkanstöße darstellen, sind im Rahmen des vom BMBF-geförderten Projekt „Fair and Good ADM“ entstanden und beschäftigten sich mit drei konkreten Beispielen für den Einsatz von ADMs in der Verwaltungspraxis:

- a. Fall 1: Risikovorhersage für Straftäter / Angeklagte durch das COMPAS-System in den USA
- b. Fall 2: Hochschulzulassung von Bewerbern auf Studienplätze in Frankreich
- c. Fall 3: Profiling von Arbeitssuchenden in Österreich durch das AMAS-System

Die Einsichten aus den explorativen Fallstudien beziehen sich dabei auf zwei übergreifende Ebenen: *Zum einen* erlauben sie Erkenntnisse über den *Policy-Making-Prozess*, der zur Entscheidung über die Einführung eines bestimmten ADM-Systems führte. Und *zum anderen* liefern sie empirische Evidenz, wie ADM-Systeme in der alltäglichen Verwaltungspraxis eingesetzt werden.



Zu beiden Themenkomplexen ist die wissenschaftliche Literatur vergleichsweise lückenhaft. Während sich in den vergangenen Jahren die Forschung zur Ausgestaltung der Algorithmen selbst und zur Frage, wie diese an ethischen Prinzipien ausgerichtet werden können, sehr dynamisch entwickelt hat (vgl. die Literatur zu „Fair Accountable and Transparent Algorithms“ (Lepri et al. 2018)) und auch regulative Vorschläge gemacht wurden (European Parliament 2019), ist sowohl die Frage der Entscheidungsfindung *über* Algorithmen als auch die Frage der konkreten Implementierung in der Verwaltungspraxis nur wenig thematisiert worden (aber, e.g. van der Voort et al. 2019). Für Deutschland liegen zudem stark rechtswissenschaftlich geprägte Studien vor (Martini et al. 2020), die zwar viele wichtige verwaltungsrechtliche Aspekte diskutieren, jedoch weniger stark auf die konkrete Umsetzung in der Verwaltungspraxis eingehen.

Auf Basis der empirischen Erkenntnisse aus den Fallstudien werden im Folgenden fünf Denkanstöße abgeleitet und begründet. Diese sind nicht als Handlungsempfehlungen zu verstehen, sondern als Anregungen, die im politischen Prozess zur Regulierung und zum Einsatz von ADM-Systemen in der öffentlichen Verwaltung aufgenommen werden können. Die folgenden Aussagen zum Einsatz von ADM-Systemen betreffen dabei ausschließlich staatliche Akteure.¹

¹¹ Manche der Vorschläge, etwa zur Notwendigkeit einer demokratisch legitimierten Prüfung der Kernelemente von ADM-Systemen, können jedoch auch die Diskussion zu einer allgemeinen Regulierung bereichern, die dann auch nicht-staatliche Akteure einschließt.

Denkanstoß 1

Über zentrale Bausteine eines ADM, wie etwa die verwendeten Kriterien oder die Optimierungsmaße in maschinellem Lernen, muss politisch in einem demokratisch legitimierten Gremium entschieden werden. Dabei ist eine gesetzliche Grundlage für jedes einzelne Politikfeld zu schaffen.

ADM-Systeme nutzen Regelsysteme, um Vorhersagen über Fälle zu machen. Diese Regeln können durch Experten vorgegeben sein – z.B. soll ein Algorithmus Bewerber*innen für einen Studiengang nach der Abiturnote, nach der Note im Fach Deutsch und nach der Zahl der Hobbies in eine Reihenfolge bringen. Die Regeln können aber auch ohne weitere Eingriffe von Experten z.B. aus korrelativen Mustern in Daten extrahiert werden (maschinelles Lernen): Wenn etwa aus den vergangenen zehn Jahren Daten zu Studierenden und ihren Leistungen vorliegen, kann ein Modell entwickelt werden, das automatisch nach Einflussfaktoren auf die Studienleistung sucht – zum Beispiel nach Zusammenhängen zwischen Studienleistungen (z.B. Abschlussnote) und interessierenden Charakteristika (z.B. Abiturnote, Note im Fach Deutsch,...). Dieses Modell kann sodann im zweiten Schritt dazu genutzt werden, (neue) Bewerber*innen für einen Studiengang in eine Reihenfolge zu bringen.

Die Fallstudien zeigen, dass es eine politische Entscheidung darüber geben muss, welche Kriterien (im zweiten Schritt) für die Vorhersage von Scores, Rankings oder Klassifizierungen durch einen Algorithmus genutzt werden. Diese Kriterien müssen diskutiert und gesetzlich verankert werden. Programmierer dürfen nicht mit der Entscheidung alleine gelassen werden, welche Variablen in ein ADM-System aufgenommen werden sollten, da dies zu demokratisch nicht legitimierten Entscheidungen und unklaren Verantwortlichkeiten führt. Im Fall der Hochschulzulassung in Frankreich über die Software APB wurde so etwa in einen Matching-Algorithmus das Kriterium der Zufallsauswahl inkludiert, um im Fall von gleich guter Qualifikation von zwei Bewerber*innen für einen Studienplatz eine automatisierte Entscheidung treffen zu können. Die Nutzung der Zufallsauswahl war gesetzlich nicht geregelt und wurde vorher auch nicht im politischen Prozess diskutiert. Diese Leerstelle führte – nach Bekanntwerden der Nutzung von Zufallsauswahl – zu gesellschaftlicher Ablehnung des gesamten Systems.

Im Fall von maschinellem Lernen müssen zudem die zur datengetriebenen Entwicklung des Modells genutzten Qualitäts- und Fairnessmaße in demokratisch legitimierten Gremien diskutiert werden. Der aus den Daten generierte Vorhersage-Algorithmus für neue Fälle (z.B. ein Ranking-Algorithmus) muss sodann öffentlich zugänglich gemacht und die für die Vorhersagen genutzten Variablen wie auch bei Experten-Systemen gesetzlich verankert werden.

In beiden Fällen ist der parlamentarische Prozess bzw. sind demokratisch über diesen Prozess legitimierte Gremien geeignete Orte für diese Entscheidungen. Zudem ist über partizipative Elemente nachzudenken, um die vom ADM-System Betroffenen Gruppen in den Prozess zu integrieren.

**Denkanstoß 2**

Die politische Entscheidung über den Einsatz eines ADM-Systems sollte immer auch die konkrete Implementation des Systems im sozialen Prozess berücksichtigen, wenn Entscheidungen erst durch die Interaktion von Mensch und Technik entstehen.

ADM-Systeme werden in der öffentlichen Verwaltung in der Regel nicht zur alleinigen Letztentscheidung sondern vielmehr in Interaktion mit Menschen zur Entscheidungsfindung eingesetzt. Wenn z.B. in den USA im Strafjustizsystem Risiko-Scores für Angeklagte zur Entscheidung über die Notwendigkeit von Untersuchungshaft für die Zeit bis zum Gerichtstermin herangezogen werden, informieren die Risiko-Scores die Entscheidung menschlicher Akteure: Sie sind ein Element von Information in Verhandlungen zwischen Richtern, Staatsanwälten und Verteidigern in der Entscheidungsfindung. Dabei ist jedoch überaus relevant, wie genau dieser Interaktionsprozess ausgestaltet ist: Die Frage, welche Person zu welchem Zeitpunkt den Output des ADM-Systems zur Risiko-Einschätzung bekommt ist zum Beispiel ein relevanter Unterschied, weil die Verfügbarkeit eines ADM-basierten Risiko-Scores als Ankerpunkt für die eigene Bewertung menschliche Entscheidungen massiv beeinflusst – so zumindest die Ergebnisse unserer Studie zum Einsatz von COMPAS in einem County im US-Staat Wisconsin.

In der Konsequenz bedeutet dies für die politische Entscheidung über die Einführung von ADM-Systemen, dass nicht nur das technische System selbst (d.h. die genutzten Kriterien und Modelle (s. Denkanstoß 1)) bewertet werden müssen, sondern vielmehr auch betrachtet werden muss, wie die konkrete Einbeziehung des ADM-Systems in das Entscheidungssystem der Verwaltung ausgestaltet ist. Möglicherweise ist es auch nötig, über einen Testbetrieb und entsprechende Begleitforschung zu prüfen, auf welche Weise sich der Entscheidungskontext von Verwaltungsmitarbeiter*innen ändert, wenn ADM-Systeme zur Verfügung stehen. Die Fallstudie zum Einsatz von Risk Scores in den USA hat zum Beispiel gezeigt, dass es ggf. zu einem übermäßigen Vertrauen in die durch das ADM-System produzierte Evidenz geben kann, wobei der „Score“ die weiteren menschlichen Überlegungen zu dominieren scheint. Daher sollte vor der politischen Entscheidung über den Einsatz eines ADM-Systems nicht nur deren technische Charakteristika (z.B. Datenqualität, Kriterien, etc.) evaluiert werden, sondern auch, wie die Systeme im konkreten Entscheidungsprozess eingebettet sind und welche Folgen die Einführung von ADMs auf menschliche Entscheidungen hat.

Denkanstoß 3

Diskurse über den Einsatz und die Ausgestaltung von ADM-Systemen müssen politisch gefördert werden, damit auch Nicht-Expertinnen die Teilnahme an einem gesellschaftlichen Diskurs über das Für und Wider von ADM-Systemen ermöglicht wird.

Die Funktionsweise von ADM-Systemen ist notwendigerweise nicht allen Menschen direkt zugänglich. Die Systeme sind komplex und in manchen Fällen ist durch laufende Anpassungen am Algorithmus selbst die nicht nachzuvollziehen, wie das System zu einem Output kommt – etwa zu einem Score, einer Klassifikation oder einem Ranking. Aus diesem Grund ist auch eine intensive gesellschaftliche Debatte über ADM-Systeme unwahrscheinlich – ein Problem, das sich auch für andere stark technisch geprägte Policies zeigt (Feindt, Schwindenhammer and Tosun 2020). Aufgrund der weitreichenden Konsequenzen, die ein Einsatz von ADM-Systemen haben kann, ist es daher notwendig, niederschwellige Möglichkeiten für den Beginn einer gesellschaftlichen und politischen Debatte zu schaffen. Die Fallstudie zum Einsatz von APB in Frankreich wie auch zum AMAS-System in Österreich hat gezeigt, dass ein politischer und gesellschaftlicher Diskurs durchaus entfacht werden kann und zu einem gesellschaftlichen Aushandlungsprozess über das Für und Wider führen kann, der schließlich auch in politische Entscheidungen mündet.

Um solche Diskurse unter Beteiligung aller gesellschaftlichen Gruppen – und nicht nur von Experten – zu ermöglichen, ist es sinnvoll, entsprechende institutionelle Vorkehrungen zu treffen. Mögliche konkrete Schritte wären etwa die Schaffung von Ombudsmännern und –frauen, die gesellschaftliche Debatten über ADM-Systeme initiieren, aufnehmen oder gestalten können. Partizipative Verfahren und die Einbindung der Zivilgesellschaft sind ebenfalls Möglichkeiten, um den Diskurs über den Einsatz von ADM-Systemen unter Einbindung aller gesellschaftlichen Gruppen zu führen.



Denkanstoß 4

Die politische Entscheidung über den Einsatz eines ADM Systems muss „rote Linien“ der Grund- und Freiheitsrechte beachten, die in manchen Anwendungsfällen aufgrund der Konsequenzen von Verwaltungsentscheidungen besonders schwer wiegen können (z.B. Strafjustiz).

In manchen Politikfeldern sind die Konsequenzen von Verwaltungsentscheidungen besonders weitreichend für die Lebenschancen und Lebensverläufe von Individuen. An vielen dieser Entscheidungen hat die demokratisch legitimierte Verfassungsordnung bereits hohe Hürden aufgestellt, die menschliche Entscheidungen nehmen müssen (z.B. im Bereich der Strafjustiz oder des Polizeirechts). Die politische Entscheidung über den Einsatz von ADM-Systemen in diesen Politikfeldern ist daher mit Blick auf die Grund- und Freiheitsrechte zu bewerten. Wenn etwa die Datengrundlage für die Entwicklung eines ADM-Systems im Bereich der Strafjustiz dadurch gekennzeichnet ist, dass erwiesenermaßen ein Großteil der „ground truth“ – also der Straftaten (und damit der Straftäter) – im Dunkelfeld verbleibt, scheint der Einsatz angesichts der weitreichenden Folgen mit Blick auf die Einschränkung von Grund- und Freiheitsrechten nicht hinnehmbar. Die demokratisch legitimierten Institutionen sollten daher „rote Linien“ definieren für den Einsatz von ADM-Systemen definieren, in denen gerade keine Abwägung zwischen Effizienz oder Qualität von Entscheidungen auf der einen Seite und Grund- und Freiheitsrechten auf der anderen Seite erfolgen kann.

über Weiterbildungsmaßnahmen

Denkanstoß 5

Nach Einführung eines ADM-Systems ist es von großer Wichtigkeit für eine nachhaltige und dauerhafte Information und Schulung der Anwender*innen durch unabhängige Expert*innen zu sorgen.

Die Fallstudien haben gezeigt, dass in manchen Fällen eine nachhaltige Schulung der Anwender*innen von ADM-Systemen in der öffentlichen Verwaltung nicht oder nicht dauerhaft erfolgt. Dies kann erhebliche Konsequenzen für die Interpretation von ADM-Outputs und damit für die menschliche Entscheidung unter Zuhilfenahme des ADM-Systems haben. Zwar wurden in allen Fällen, die im Rahmen des Projekts untersucht wurden, die Anwender im Rahmen der Einführung des ADM-Systems informiert und geschult. Jedoch lag kein Prozess einer dauerhaften Begleitung durch Trainings vor. Dies ist dann besonders relevant, wenn die Mitarbeiter*innen in der Verwaltung fluktuieren oder wenn sich das ADM-System über die Zeit verändert. Daher sollte sichergestellt werden, dass Anwender*innen dauerhaft geschult werden.

Zudem halten wir es für zentral, dass die Schulungen entweder durch unabhängige Experten erfolgt oder diese zumindest den Schulungsprozess und die -inhalte regelmäßig überprüfen. Da ADM-Systeme als Produkte auch den Kriterien der freien Wirtschaft und damit des Marktes unterliegen, ist es wichtig, dass Schulungen nicht von Vermarktungsaspekten dominiert werden, sondern tatsächlich ein wissenschaftlich fundiertes Bild der Vor- und Nachteile vermitteln.

Denkanstoß 6

ADM-Systeme sind nicht alternativlos. Auch ohne die Nutzung von Softwaresystemen kann die Analyse von „big data“ die Informationsbasis für menschliche Entscheidungsträger*innen verbessern.

Die Fallstudien haben gezeigt, dass die Nutzung empirischer Evidenz und die Aufbereitung dieser Evidenz für Entscheidungsträger*innen die Informationsgrundlage stark verbessern kann. So sind etwa die Informationen über die beruflichen Karrieren von Arbeitssuchenden wichtige empirische Daten, aus denen Hinweise zu möglichen Einflussfaktoren auf die Chancen der Arbeitsmarktintegration abgeleitet werden können. Diese werden in der wissenschaftlichen Forschung auch entsprechend genutzt, um Einflussfaktoren auf Erwerbsverläufe zu analysieren.

Dass diese Informationen in Form eines ADM-Systems aufbereitet und über ein Softwareprogramm für Benutzer*innen zur Verfügung gestellt werden, ist eine mögliche, aber nicht eine notwendige Folge. Wenn etwa eine Untersuchung der Mensch-Technik-Interaktion zeigt, dass eine solche Implementation des ADM-Systems im Entscheidungsprozess nicht zielführend ist (siehe Denkanstoß 2), bedeutet dies nicht, auf die Ergebnisse empirischer Evidenz zu verzichten. Vielmehr sollten im Sinne einer evidenzbasierten Entscheidung alle wissenschaftlichen Erkenntnisse in die menschliche Entscheidungsfindung einfließen. Eine Entscheidung gegen die Nutzung eines ADM-Systems ist daher nicht gleichbedeutend mit einer uninformierten Entscheidung. Vielmehr ist es notwendig, über alternative Formate nachzudenken, wie Entscheidungsträger*innen in der öffentlichen Verwaltung – etwa in der Arbeitsvermittlung – möglichst gut informierte Entscheidungen auf Basis der verfügbaren empirischen Evidenz treffen können.

Um empirische Erkenntnisse für die Verwaltungspraxis nutzbar zu machen, sind dabei zwei Elemente zentral: Zum einen gilt – wie auch bei der Nutzung von ADM-Systemen (siehe Denkanstoß 5) – dass Mitarbeiter*innen über Weiterbildungsmaßnahmen geschult werden müssen und dabei systematisch über den Stand der wissenschaftlichen Forschung zu ihrem Arbeitsfeld informiert werden. Zum anderen ist es notwendig, die Rückkopplung zwischen dem Outcome einer Entscheidung und der Entscheiderin selbst zu stärken. Zu selten wissen Verwaltungsangestellte, ob eine bestimmte Entscheidung (etwa über ein Trainingsprogramm für Arbeitssuchende) am Ende erfolgreich war. Wenn jedoch dieses Feedback fehlt, sind die Anreize für Verwaltungsmitarbeiter*innen eigene Entscheidungen zu hinterfragen (oder auch zu bestätigen) gering – und damit geht eine wichtige Information verloren, die zukünftige Entscheidungen verbessern könnte.

Literatur



- Egbert, S. and S. Krasmann (2019) Predictive Policing. Eine ethnographische Studie neuer Technologien zur Vorhersage von Straftaten und ihre Folgen für die polizeiliche Praxis. Projektabschlussbericht. Hamburg: Universität Hamburg.
- European Parliament (2019) A governance framework for algorithmic accountability and transparency. Brussels: European Parliamentary Research Service, Scientific Foresight Unit.
- Feindt, P. H., S. Schwindenhammer and J. Tosun (2020) Politicization, Depoliticization and Policy Change: A Comparative Theoretical Perspective on Agri-food Policy, *Journal of Comparative Policy Analysis: Research and Practice*, 1-17.
- Lepri, B., N. Oliver, E. Letouzé, A. Pentland and P. Vinck (2018) Fair, Transparent, and Accountable Algorithmic Decision-making Processes, *Philosophy & Technology*, 31, 4, 611-27.
- Martini, M. (2019) *Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz*. Berlin, Heidelberg: Springer.
- Martini, M., J. Botta, D. Nink and M. Kolain (2020) *Automatisch erlaubt? Fünf Anwendungsfälle algorithmischer Systeme auf dem juristischen Prüfstand*. Gütersloh: Bertelsmann Stiftung.
- van der Voort, H. G., A. J. Klievink, M. Arnaboldi and A. J. Meijer (2019) Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making?, *Government Information Quarterly*, 36, 1, 27-38.
- Wissenschaftliche Dienste des Bundestages (2019) *Einsatz künstlicher Intelligenz in öffentlichen Verwaltungen ausgewählter Länder*. Sachstand WD 10 - 3000 - 037/19. Berlin: Deutscher Bundestag.